

# LA EVALUACIÓN POR PARES Y LA PUBLICACIÓN: LECCIONES PARA ABOGADOS\*

Susan Haack

*University of Miami*

**RESUMEN.** La Corte Suprema estadounidense en el caso *Daubert v. Merrell Dow Pharmaceuticals Inc.* especificó como un posible criterio a tomar en consideración por los jueces para decidir la admisibilidad de las pruebas científicas la revisión por pares y la publicación de trabajos sobre la teoría o técnica subyacente al elemento de prueba en cuestión. En este trabajo se abordan algunos de los problemas que los actuales sistemas de revisión por pares y publicación enfrentan y que permiten cuestionar la utilidad del criterio si el objetivo de los tribunales es valorar la calidad de la información científica.

**Palabras clave:** Evaluación por pares, publicaciones, pruebas científicas, fiabilidad de la prueba pericial.

## Peer Review and Publication: Lessons for Lawyers

**ABSTRACT:** In *Daubert v. Merrell Dow Pharmaceuticals Inc.*, the Supreme Court of the United States said that a pertinent consideration in order to determine the admissibility of scientific evidence is whether the theory or technique has been subjected to peer review and publication. This work analyses some of the problems faced by the current systems of peer review and publication, which also call into question their use as an indicia of the reliability of scientific information.

**Keywords:** Peer review, publications, scientific evidence, reliability of expert evidence.

---

\* Fecha de recepción: 20 de noviembre de 2013. Fecha de aceptación: 10 de diciembre de 2013. Este trabajo fue publicado en inglés en *Stetson Law Review*, 36, núm. 3 (2007), 789-819.

Una consideración pertinente [para determinar si una teoría o una técnica es conocimiento científico que asistirá al juzgador de los hechos] es si la teoría o técnica ha sido sometida a una evaluación por pares y publicada.

*Daubert v. Merrell Dow Pharmaceuticals, Inc.* (1993)<sup>1</sup>

La frase «evaluación por pares» connota la revisión del trabajo científico o de otro corte académico por terceros que se supone tienen *expertise* en el área relevante («pares»). Específicamente, y más relevante para el propósito de este artículo, se refiere a la evaluación de manuscritos con el objetivo de determinar qué trabajos serán publicados en las revistas especializadas y cuáles libros serán publicados por editoriales académicas (contexto en el cual también es conocido como «arbitraje», «dictamen de pares», «proceso de revisión entre pares para la publicación»)<sup>2</sup>. Ocasionalmente, sin embargo, la frase se llega a usar en un sentido mucho más amplio para referirse a todo el largo camino de escrutinios a los que es sometido un trabajo científico en una comunidad científica y los esfuerzos de otros para usarlo como punto de partida<sup>3</sup>, un proceso largo en el que la evaluación por pares en sentido estricto es sólo una pequeña parte.

Estas dos concepciones de la evaluación por pares, la amplia y la estricta, entran en juego en los argumentos sobre la admisibilidad de las pruebas periciales presentadas por los demandantes en el caso *Daubert*<sup>4</sup>. En 1989, concediendo un fallo sumario a Merrell Dow con el fundamento de que las pruebas ofrecidas por los Daubert sobre la relación causal eran inadmisibles, la *District Court* había enfatizado que «ninguno de los estudios *publicados* mostraba una correlación estadísticamente significativa entre el uso del Bendectin y los defectos congénitos»<sup>5</sup>; y, en 1991, el juez KOZINSKI al confirmar esta decisión, argumentando que «ningún estudio epidemiológico publicado ha establecido una correlación estadísticamente significativa entre el Bendectin y los defectos congénitos» y que «el proceso normal de evaluación por pares [...] es uno de los rasgos distintivos de la investigación científica fiable», también consideró que las publicaciones arbitradas eran un factor clave<sup>6</sup>.

En 1993, cuando el caso llegó a la Corte Suprema, un *amicus brief*\* presentado por CHUBIN *et al.*, criticó la confianza de los tribunales inferiores en las publicacio-

<sup>1</sup> *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579, 593 (1993) [en adelante *Daubert* (1993)].

<sup>2</sup> En inglés la frase «*peer review*» algunas veces también hace referencia a la evaluación de la clínica realizada por los practicantes *senior* en un área de conocimiento (contexto en el cual se conoce como «*clinical peer review*»); a la evaluación realizada para revisar propuestas para subvenciones propuestas y decidir qué proyectos financiar (contexto en el que se llama «*grant peer review*» o «*merit review*»); y a la evaluación de resúmenes o artículos sometidos para determinar cuáles serán presentados en un congreso.

<sup>3</sup> «La evaluación por pares en su sentido más amplio puede decirse que ha existido desde que la gente empezó a identificar y comunicar lo que consideraban como nuevo conocimiento» (KRONICK, 1990: 1321).

<sup>4</sup> CHAN, 1995: 113.

<sup>5</sup> *Daubert v. Merrell Dow Pharms., Inc.*, 727 F.Supp. 570, 575 (S.D. Cal. 1989) [en adelante *Daubert* (1989)].

<sup>6</sup> *Daubert v. Merrell Dow Pharms., Inc.*, 951 F.2d 1128, 1129, 1131 n.3 (9th Cir. 1991) [en adelante *Daubert* (1991)].

\* N. de la T.: un *amicus brief* es un informe (*brief*) presentado por un tercero (*amicus*) que tiene un fuerte interés u opinión en el tema del litigio del que *no* es parte. Dicho informe puede ser solicitado por el propio tribunal u ofrecido de forma voluntaria por el *amicus*, en este último caso tiene que contar con la anuencia

nes arbitradas, argumentando que «el sistema de evaluación por pares está diseñado para proveer un punto de inicio común y conveniente para el debate científico, más no la última palabra del conocimiento científico existente» y que «contrario al mito “generalmente aceptado”, la publicación de un artículo en una revista con *referato* no asegura que la investigación, las metodologías [o] los análisis [...] sean verdaderos, exactos, [...] fiables o ciertos, o indicadores de “buena ciencia”»<sup>7</sup>. Y mientras que la decisión del magistrado BLACKMUN incluía «la revisión por pares y la publicación» como un factor que los tribunales podrían considerar para determinar si una prueba científica es «fiable» en el sentido exigido para ser admisible, lo hacía de forma muy restrictiva y cautelosa —reconociendo que el proceso de evaluación por pares para la publicación no garantiza la «fiabilidad probatoria», pues bien podrían frenarse trabajos bien fundados porque son innovadores; y que un mejor indicador es sobrevivir a largo plazo al escrutinio de la comunidad científica, *i. e.*, a la revisión por pares en un sentido amplio—<sup>8</sup>.

Finalmente, en 1995, una vez que la Corte Suprema había ordenado la reposición y traslado del caso, el juez KOZINSKI reconociendo, como lo había hecho también el magistrado BLACKMUN, que una publicación arbitrada no garantizaba la fiabilidad del testimonio, argumentó que no obstante el hecho «[d]e que una investigación sea aceptada para su publicación en una revista científica de buena reputación [...] es una indicación significativa de que [...] ésta satisface al menos los criterios mínimos de la buena ciencia»<sup>9</sup>. Así pues, dado que «[n]inguno de los peritos de los demandantes ha publicado en una revista científica su investigación sobre el Bendectin», confirmó una vez más el fallo sumario decidido por el tribunal inferior<sup>10</sup>.

El objetivo de este trabajo es entender cómo funciona el proceso de evaluación por pares, cuán buen indicador es de que una prueba científica es «fiable» en el sentido jurídicamente relevante y cómo los tribunales estadounidenses podrían hacer un mejor uso de este factor *Daubert*. Así, la mayor parte del resto del artículo me centraré en la evaluación por pares en sentido estricto —es decir, en el proceso de revisión por pares a efectos de la publicación—. Iniciaré con un bosquejo sobre los orígenes de esta práctica, el desigual proceso mediante el cual gradualmente se fue estandarizando en las revistas científicas y médicas y los muchos roles que ahora juega (sección 1); el siguiente paso será articular la explicación de por qué es útil el proceso de evaluación por pares y las limitaciones inherentes del sistema como un mecanismo de control de calidad (sección 2); y luego, hacer un sondeo sobre los cambios en la ciencia, en las publicaciones científicas y en la academia que han puesto bajo severa prueba al sistema de evaluación por pares y sobre algunos ejemplos recientes en los

---

del tribunal. En latín también es conocido como *Amicus Curiae*, es decir, amigo del tribunal. *Vid. Black's Law Dictionary*, 1990, 83.

<sup>7</sup> Informe presentado como *amicus brief* de D. E. CHUBIN *et al.*, en apoyo del demandante, *Daubert*, 1993 (*supra* nota 1) (núm. 92-102), 1992 WL 12006443, \*8, \*13 («Informe de Daryl Chubin *et al.*») (itálicas y cita omitida).

<sup>8</sup> *Daubert* (1993), 509 U.S., 593.

<sup>9</sup> *Daubert v. Merrell Dow Pharms., Inc.*, 43 F.3d 1311, 1318 (9th Cir. 1995) [en adelante, *Daubert* (1995)]. Para un análisis más detallado de la sentencia del juez KOZINSKI en este caso, *vid.* HAACK (2008), reimpresso en HAACK (2014: 185-88).

<sup>10</sup> *Ibid.*, 1318, 1332.

que un trabajo poco sólido o incluso fraudulento ha pasado la evaluación por pares (sección 3).

Sin embargo, en la siguiente sección, cuando se analicen las observaciones del magistrado BLACKMUN sobre «la evaluación por pares y la publicación» en la decisión del caso *Daubert*, veremos que la «evaluación por pares» en el sentido amplio también tendrá, junto a su sentido estricto, cierto rol. Dicho de forma somera, defenderé que no es de mucha ayuda práctica la idea planteada por el magistrado BLACKMUN sobre que las publicaciones arbitradas no son condición necesaria ni suficiente de la fiabilidad probatoria y que un mejor indicador de la validez científica es sobrevivir a largo plazo al proceso de revisión de la comunidad científica; tampoco es de ayuda la sugerencia que dio a los tribunales, consistente en que las publicaciones arbitradas pueden ser un factor relevante pero no determinante para decidir la admisibilidad (sección 4).

Posteriormente, si una prueba pericial ha sido excluida en parte porque no está fundada en una publicación arbitrada o si ha sido admitida a pesar de no tener tales fundamentos, los tribunales, por lo general, parecen no haberse preguntado cuestiones que podrían arrojar luz sobre qué significa que una concreta publicación está arbitrada, o que no lo está. Pero, como veremos, el escrutinio de inusual sentido común realizado por un tribunal de Pensilvania sobre la literatura arbitrada del Bendectin muestra cuán débil puede ser la «evaluación por pares y la publicación» —y deja a uno incómodamente preguntándose sobre la forma en que este «factor *Daubert*» fue objeto de atención de los tribunales por primera vez (sección 5)—<sup>11</sup>.

## 1. EL PROCESO DE EVALUACIÓN POR PARES PARA LA PUBLICACIÓN: SUS RAÍCES HISTÓRICAS Y SUS ROLES ACTUALES

A los científicos siempre les ha interesado que su trabajo sea reconocido como suyo y se han preocupado sobre lo que de forma encantadora R. BOYLE describió como «*philosophicall robbery*» («robo filosófico»), también conocido como plagio<sup>12</sup>. Incluso antes de que las *Philosophical Transactions* de la *Royal Society of London* fueran creadas en 1655, la Sociedad quiso darle un toque oficial a la prioridad del científico en el descubrimiento mediante el recordatorio de la fecha en la que recibió una carta anunciando un experimento o una observación<sup>13</sup>. Como H. OLDENBURG, el primer editor de las *Transactions*, le dijo a BOYLE, la Sociedad sería: «Muy cuidadosa en registrar tanto a la persona y el momento de cualquier nueva cuestión comunicada, como los resultados mismos; a través de lo cual el honor de un invento será preservado inviolable para toda la posteridad»<sup>14</sup>. Gradualmente, las *Transactions* empezaron a indicar qué trabajos habían sido evaluados por representantes de la Sociedad antes de su publicación y cuáles no lo habían sido; y, en 1702 el *Journal des Scavans*, fundado

<sup>11</sup> *Blum v. Merrell Dow Pharms., Inc.*, 764 A.2d 1, 3 (Pa. 2000).

<sup>12</sup> ZUCKERMAN y MERTON, 1971: 66.

<sup>13</sup> *Ibid.*

<sup>14</sup> *Ibid.*, 70 [citando *The Correspondence of Henry Oldenburg*, vol. 1, 319 (A. R. HALL y M. B. HALL —eds. y trads.—, Madison, University of Wisconsin Press, 1966)].

justo antes de *Transactions*, había asignado a varios miembros del consejo editorial la responsabilidad de filtrar los trabajos sometidos en un área determinada<sup>15</sup>.

En el curso del siglo XVIII muchas otras importantes publicaciones en el área médica y científica adoptaron lo que ahora llamaríamos «evaluación por pares»: en 1731, el Prefacio al primer volumen de los *Medical Essays and Observations*, publicados por la *Royal Society of Edinburgh*, anunció que las memorias enviadas por correspondencia eran distribuidas atendiendo al tema entre aquellos miembros más versados en él<sup>16</sup>; en 1752, la *Royal Society* estableció un comité autorizado para llamar a «cualquier otro miembro de la Sociedad que tuviera conocimiento y buenas habilidades en una rama particular de la Ciencia de la que fuera objeto de estudio» para revisar un artículo sometido a *Transactions*<sup>17</sup>; en 1782, la regulación de la *Académie Royale de Médecine* estableció que «nada sería impreso en *Histoire* o en el *Receuil des memoires* de la Sociedad [...] sin que una asamblea especialmente convocada para ese propósito hubiese decidido por el voto de la mayoría su publicación»<sup>18</sup>; y en 1785, la *Literary and Philosophical Society of Manchester* estableció un comité revisor para seleccionar los trabajos «con la mayor imparcialidad y la estricta atención a sus méritos comparativos» tanto como fuese posible<sup>19</sup>.

De acuerdo con el historiador J. BURNHAM, la extensión y evolución de la práctica del proceso de evaluación por pares entre el siglo XIX y las primeras décadas del XX no fue sistemática ni metódica<sup>20</sup>. Algunas de las primeras revistas médicas de inicios del siglo XIX fueron, como BURNHAM apuntó, «vehículos personales» para editores como Th. WAKELY, fundador de *The Lancet*, o H. MAUNSELL, uno de los fundadores de la *Dublin Medical Press*, que posteriormente también se convirtió en dueño del *Dublin Evening Mail*<sup>21</sup>. Algo muy cercano a la situación actual, las revistas científicas y médicas fueron publicaciones oficiales de institutos europeos de investigación (especialmente alemanes); estos foros más especializados confiaban en el juicio experto del editor o de los colegas que componían la planilla editorial, pero publicaban principalmente los trabajos de los miembros de los institutos<sup>22</sup>.

Pero en los primeros días de las publicaciones tanto científicas como médicas, el principal problema de un editor era encontrar material suficiente para llenar sus páginas, y no decidir qué trabajo publicar entre los muchos trabajos presentados. En 1876 un comentador observó que «exigir artículos breves y reportes de casos singulares deja muchas páginas disponibles en blanco»<sup>23</sup>. Las publicaciones de periodicidad semanal y mensual son omnívoros e insaciables en su solicitud de contribuciones<sup>24</sup>; incluso en

<sup>15</sup> LOCK, 1985: 2.

<sup>16</sup> KRONICK, 1990: 1321 [citando I *Essais et Observations de Médecine de la Société D'edimbourg*, vol. 1, preface (1740)].

<sup>17</sup> *Ibid.*, 1321 (citando a ZIMAN, 1968: 318).

<sup>18</sup> *Ibid.*, 1321-1322 [citando 13 *Histoire De L'académie Royale De Médecine* 19-21 (1782)].

<sup>19</sup> *Ibid.*, 1322 [citando 1 *Memoirs of The Literary and Philosophical Society of Manchester*, prefacio (1785)].

<sup>20</sup> BURNHAM, 1990: 1327-1328.

<sup>21</sup> *Ibid.*, 1324.

<sup>22</sup> *Ibid.*, 1324.

<sup>23</sup> *Ibid.*, 1325.

<sup>24</sup> *Ibid.*, 1325 [citando a BILLINGS (1876), *Literature and Institutions, American Journal of Medical Science*, 72, 439, 460].

1921 el editor de la *Journal of Neurology and Psychopathology* se quejó con un corresponsal de la dificultad de obtener material suficiente para su revista<sup>25</sup>. Fue solamente hasta después de la Segunda Guerra Mundial que la evaluación por pares, tal como la conocemos hoy, se convirtió en una práctica común de las revistas médicas y científicas<sup>26</sup>; entonces, un cambio significativo en el número de artículos ofrecidos significó que los editores ya no estuvieran buscando material para rellenar sus páginas, sino una forma de seleccionar los trabajos que se publicarían<sup>27</sup>.

Hoy en día, el proceso de evaluación por pares para publicar es normal en las revistas médicas y científicas<sup>28</sup>; y también un procedimiento estándar en las publicaciones académicas de otras áreas, incluidas las humanidades (aunque *no* en las revistas jurídicas estadounidenses)<sup>29</sup>. Esto, en consecuencia, también se convirtió en un factor muy importante para la economía de publicaciones científicas, médicas y otras publicaciones académicas; porque el prestigio de grandes casas editoriales tanto científicas como médicas y de las editoriales académicas y, con ello, los altos precios que pueden establecer para sus publicaciones<sup>30</sup>, deriva en parte de que tales publicaciones son percibidas como «certificadas» dada la evaluación por pares.

Por otra parte, la evaluación por pares está ahora muy arraigada en los sistemas de promoción y de permanencia de los profesores en las universidades, que exigen publicaciones arbitradas o son menos favorables hacia publicaciones que no son arbitradas y, además, llegan a considerar que la participación de uno de los miembros de la facultad como *réferi* en revistas académicas o editoriales es parte de su «servicio»<sup>31</sup>. De hecho, a menudo las universidades usan el proceso de evaluación por pares como sustituto —es muy tentador decir, como un sustituto ocioso— de una evaluación sustantiva de la calidad del trabajo de una persona. Tal como en una editorial inusualmente franca de *Nature* se manifestó la queja: «Las universidades [...] han caído en el desprolijo

<sup>25</sup> *Ibid.*, 1325 [citando una carta de C. Stanford Read a Smith Ely Jelliffe (con fecha 3 de febrero de 1921), en *Papers of Smith Ely Jelliffe*, 1866-1940 (la carta original está en la Library of Congress, Washington, D. C., Box 16)].

<sup>26</sup> SUN, 1989: 910. James McKeen CATTELL, quien editó la revista *Science* desde 1894 hasta su muerte en 1945, aparentemente confiaba ampliamente en su hijo (quien tenía un título de Harvard) para ayudar a revisar los artículos propuestos; pero cuando la *American Association for the Advancement of Science* asumió la revista en 1945, se instituyó un sistema de evaluación por pares.

<sup>27</sup> BURNHAM, 1990: 1236-1237.

<sup>28</sup> LOCK, 1985: 3. En 1985, al menos tres cuartos de las revistas científicas más importantes en occidente, confiaban en la evaluación por pares. En la página siguiente, LOCK (1985: 4) reportó que en 1980 las 100 revistas científicas soviéticas también usaban la evaluación por pares [citaba a A. RELMAN, «Moscow in January», *New England Journal of Medicine*, 302 (1980): 523].

<sup>29</sup> En las revistas jurídicas son usualmente los estudiantes, no los académicos, quienes deciden qué artículos son aceptados. Al respecto, *vid.* POSNER, 2004: 57 (señalando que son los estudiantes editores quienes en última instancia deciden qué artículos publicar y argumenta que las publicaciones en revistas jurídicas son, por esa razón, menos rigurosamente controladas que las publicaciones en otras áreas académicas). Pero, *vid.* el informe presentado como *amicus brief* de D. CHUBIN *et al.* (*supra* nota 7), \*8 n. 8 (señalando que las revistas jurídicas son en algunos aspectos *más* rigurosas, dado que los estudiantes editores, que revisan cada cita y nota al pie, pasan mucho más tiempo en los artículos que aquel que pueden pasar los evaluadores pares de las revistas científicas).

<sup>30</sup> En octubre de 2003, científicos de la Universidad de California, San Francisco, organizaron una protesta en contra de Elsevier por una factura de 91.000 dólares para seis revistas de biología; finalmente, la universidad negoció «una reducción del 25% del precio a \$7,7 millones al año para 1.200 Elsevier publicaciones» (WYSOCKI, 2005: A1, A8).

<sup>31</sup> EDITORIAL, 1983: 361.



hábito de sustituir su criterio sobre sus propios logros por el juicio de evaluadores externos, ofrecido por el subgrupo apropiado del sistema de evaluación por pares»<sup>32</sup>.

Tal como alguna vez P. BRIDGMAN observó, mientras «[u]n perro está contento con dar tres vueltas antes de echarse», un ser humano tiene que pensar en alguna razón sobre por qué ésta es la *mejor* forma de echarse; «[n]o hay una sola institución humana que no ha tenido su origen en el éxito o en el fracaso de una manera más o menos arbitraria, sin embargo, cada una de estas instituciones está justificada por un argumento que pretende mostrar que racionalmente es la mejor posible»<sup>33</sup>. Entonces no es sorprendente que, así como el proceso de evaluación por pares se ha extendido y afianzado en las publicaciones académicas y en la academia misma, algunos se inclinan a exagerar sus virtudes —pensando en el sistema, no como un basto filtro preliminar, sino como un muy buen indicador de calidad—. En 1968, J. ZIMAN describe al *referato* como «el soporte central del que depende toda la empresa científica»; y, más recientemente, P. GROSS, investigador en las ciencias de la vida, escribe que «arbitrada» es —hablando «vagamente pero no incorrectamente»— un «tipo de antónimo» de «sesgo»<sup>34</sup>.

Pero incluso si el sistema de evaluación por pares para publicar trabajara perfectamente, estaría inherentemente limitado en aquello que podría hacer para asegurar la calidad —de la cual, la «fiabilidad» en el sentido jurídicamente relevante es sólo una dimensión—; es más, hay buenas razones para temer que, dados los cambios que ha habido en la escala y la cultura de las ciencias desde que el sistema se estandarizó, actualmente el sistema trabaja de forma muy imperfecta.

## 2. EL PROCESO DE EVALUACIÓN POR PARES PARA LA PUBLICACIÓN: SU UTILIDAD Y SUS LIMITACIONES INHERENTES

En 1946, justo cuando esta práctica se convirtió en un proceso estándar de las revistas científicas, M. POLANYI pronunció su clásica declaración sobre la utilidad epistemológica del proceso de evaluación por pares. Es necesario un sistema para racionar las limitadas oportunidades de publicación, argumentó; dado que la empresa científica depende de compartir efectivamente las pruebas y del escrutinio mutuo, sin un sistema como éste los científicos se verán obligados a perder su tiempo filtrando el trabajo de excéntricos e incompetentes para encontrar lo que vale la pena:

Supongamos [...] que no se impusiera ninguna limitación evaluativa para la publicación de contribuciones científicas en las revistas. La selección —que es indispensable en vista del limitado espacio— tendría que hacerse mediante algún método neutral —digamos eligiendo al azar. Inmediatamente las revistas estarían inundadas de basura y los trabajos valiosos serían desplazados. Hay siempre muchos excéntricos que envían cantidades de basura sin sentido. Se difundiría material deficiente, confuso, fantasioso, ordinario, no pertinente. Los estafadores y los ineptos combinando todas las variantes de engaños y autoengaños busca-

<sup>32</sup> EDITORIAL, 1983: 361.

<sup>33</sup> BRIDGMAN, 1955: 368.

<sup>34</sup> Correo electrónico de P. R. GROSS, profesor emérito de las ciencias de la vida en la Universidad de Virginia, a S. HAACK, el 11 de julio del 2005 (copia en archivo de la autora). *Vid.* también ZIMAN (1968: 111).

rían publicidad. Enterradas entre tanto material engañoso, las pocas publicaciones valiosas restantes apenas tendrían alguna oportunidad de ser reconocidas<sup>35</sup>.

Continúa POLANYI, racionar mediante el proceso de revisión por pares es una forma de asegurar que aquello que es publicado al menos cumple con los estándares mínimos de competencia profesional.

Ninguna contribución propuesta a la ciencia tiene oportunidad de ser generalmente conocida sino es publicada en prensa; y sus oportunidades de reconocimiento son muy pocas a menos que sea publicada en una de las revistas científicas más importantes. Los *réferis* y los editores de tales revistas son responsables de excluir todo aquel material que consideren poco sólido o no pertinente. Ellos tienen la responsabilidad de proteger un estándar mínimo para toda la literatura científica publicada<sup>36</sup>.

Las frases clave para nuestros propósitos son «poco sólido o no pertinente» y «proteger un estándar mínimo»<sup>37</sup>.

«Poco sólido» y «estándar mínimo» sugieren que no se puede esperar que el proceso de revisión por pares garantice la verdad, la validez de la metodología o la rigurosidad de los cálculos estadísticos usados, etc. Desde sus orígenes, los editores científicos han enfatizado que ellos y sus evaluadores no tienen otra opción que confiar en la integridad de los autores. En 1665 Denis DE SALLO, el primer editor del *Journal des Scavans*, escribió en el primer número que «[su] objetivo sería dar a conocer las ideas de otros sin garantizarlas»<sup>38</sup>; la declaración de 1731 sobre la política del *referato* de la *Edinburgh Society* concluye con la observación de que «la responsabilidad sobre la verdad de los hechos, la solidez del razonamiento y la precisión de los cálculos está totalmente deslindada; y debe descansar únicamente en el conocimiento, el juicio o la habilidad de los autores, quienes respetuosamente han suministrado dicha información»<sup>39</sup>.

Y el hecho de que POLANYI usara la palabra «no pertinente» nos recuerda que los editores y los *réferis* no están interesados *solamente* en la verdad, la solidez metodológica, etc.; también les importa, razonablemente, el interés de los trabajos, la legibilidad de los artículos y su idoneidad para una revista particular<sup>40</sup>. Como dijo el anterior editor de *Journal of the National Cancer Institute*, al escribir sobre «[f]iabilidad [...] y otros objetivos inapropiados de la evaluación por pares», «las decisiones editoriales pueden hacer, hacen y deben hacer uso de otros criterios [...] [tales como] la originalidad, la idoneidad del tema dada la revista, [...] la necesidad de un balance de temas en las revistas que tienen una cobertura amplia, la importancia de los descubrimientos para los lectores»<sup>41</sup>.

POLANYI deja claro que no es la publicación arbitrada en sí misma lo que da cierta autoridad a los resultados científicos, sino lo que pasa *después* de la publicación del trabajo:

<sup>35</sup> POLANYI, 1946: 35-36.

<sup>36</sup> *Ibid.*, 33.

<sup>37</sup> *Ibid.*, 33.

<sup>38</sup> RENNIE, 2003: 2.

<sup>39</sup> KRONICK, 1990: 1322.

<sup>40</sup> POLANYI, 1946: 33-34.

<sup>41</sup> BAILAR, 1991: 137-138.



Mediante su publicación, un trabajo es expuesto al escrutinio abierto de todos los científicos que se formaran una opinión sobre su valía y, posiblemente, también la expresaran. Pueden dudar o rechazar todas las afirmaciones ahí vertidas, en tanto que, el autor probablemente las defenderá. Después de un tiempo, una opinión más o menos estable prevalecerá. La tercera etapa del escrutinio público por la que debe pasar una contribución a la ciencia para ser generalmente conocida y establecida es su incorporación en los libros de texto o, al menos, en los libros estándar de referencia<sup>42</sup>.

Por otra parte, también reconoce que el sistema de evaluación por pares, incluso en su tarea más modesta de «proteger un estándar mínimo», sólo sería exitoso bajo ciertos presupuestos:

Si todo científico cada mañana se pusiera a trabajar con la intención de hacer la mejor charlatanería que, sin ser descubierto, le ayudara a tener un buen puesto, pronto no existirían estándares efectivos con los que se pudiera detectar tal engaño [...]. Sólo si los científicos permanecen leales a los ideales científicos en lugar de tratar de lograr el éxito con sus colegas, éstos pueden formar una comunidad que defienda tales ideales<sup>43</sup>.

Obviamente (aunque POLANYI no lo dice explícitamente) la efectividad del sistema no sólo depende de la integridad de los autores sino también de la integridad de los evaluadores, editores y editoriales. Y el problema no es sólo que el sistema falle si cada científico trabaja haciendo «la mejor charlatanería» con la que puedan salirse con la suya; sino también que funcionará de forma menos efectiva entre más pesadas sean las cargas de trabajo sobre los revisores y los editores, entre más fuertes sean las presiones sobre las revistas y las tentaciones para que los científicos tomen ciertos atajos, sean estratégicamente vagos o incluso inventen resultados.

### 3. EL PROCESO DE REVISIÓN POR PARES PARA LA PUBLICACIÓN: RECIENTES PRESIONES Y TENSIONES, DEBILIDADES Y FRACASOS

Incluso en circunstancias ideales, los evaluadores están mejor situados para juzgar la legibilidad de un trabajo o el interés del tema o sus resultados que su veracidad o precisión y pueden con buena fe rechazar trabajos importantes que son demasiado innovadores como para parecer plausibles; así pues, quizá no resulta sorprendente que en 1994 el historiador de la ciencia H. F. JUDSON, describiendo las «transformaciones estructurales» que tienen lugar en las ciencias, tuvo que incluir en su listado los «estándares siempre más bajos y la tendencia inherente hacia la corrupción del proceso de evaluación por pares y el *referato*»<sup>44</sup>. Actualmente hay muchas presiones que ponen al sistema de revisión por pares bajo severas tensiones: la explosión de las publicaciones científicas y médicas; la creciente influencia financiera de las grandes compañías farmacéuticas sobre las revistas médicas; la presión sobre los científicos jóvenes para obtener fondos y para publicar; la tentación de buscar convertirse en una celebridad; las siempre más grandes oportunidades para ganar dinero como perrito, etcétera<sup>45</sup>.

<sup>42</sup> POLANYI, 1946: 33-34.

<sup>43</sup> *Ibid.*, 40.

<sup>44</sup> JUDSON, 1994: 92.

<sup>45</sup> HAACK, 2003: 27-29, 107-109.

Hay diferencias entre las revistas científicas y las médicas, pero, en términos generales el proceso de *referato* funciona así: un editor hace lo que LOCK describe como «priorización»: «Clasificando los artículos en obras maestras innegables, basura inquestionable y el resto que precisa de cuidadosa consideración»<sup>46</sup>; para ello —la gran mayoría de los editores— eligen uno o dos (raramente más) *réferis* que revisan cada trabajo seleccionado, generalmente informando a éstos el nombre del autor pero no viceversa<sup>47</sup>; habitualmente se les da a los *réferis* una lista de criterios para que revisen diversos aspectos de estilo, presentación y ciertos tipos de error<sup>48</sup>; los evaluadores tienen un tiempo límite, frecuentemente de no más de dos semanas, para dar sus evaluaciones y recomendaciones<sup>49</sup>; y éstos pasan un promedio de 2,4 horas evaluando un manuscrito que normalmente no requiere un simple veredicto de «sí o no», sino hacer sugerencias sobre cómo el trabajo en cuestión puede ser mejorado<sup>50</sup>. Muchas revistas no revisan el cálculo estadístico de los trabajos que son aceptados<sup>51</sup>; y los evaluadores no están en posición de repetir los experimentos o los estudios de los autores, pues tendrían que invertir mucho dinero y/o tiempo. Las tasas de aceptación varían ampliamente de área en área; en aquellas en que la tasa es baja, la mayoría de los trabajos inicialmente propuestos, pero que han sido rechazados por una o más de las revistas más solicitadas, finalmente aparecen en alguna publicación de menor rango y así un trabajo «puede ser rechazado de diez a veinte veces antes de que finalmente sea aceptado»<sup>52</sup>. Los capítulos de libros son usualmente invitados, no sujetos a la evaluación por pares.

<sup>46</sup> LOCK, 1985: 6.

<sup>47</sup> Algunas revistas están yendo hacia una evaluación «abierta», en la que los autores también conocen los nombres de los evaluadores. Vid. SMITH, 1997: 760 (argumentando que la evaluación abierta es la forma más ética porque pone a los autores y a los evaluadores en posiciones iguales y permite mejorar la rendición de cuentas). En contraste, en las revistas de filosofía, y hasta donde sé en las revistas de la humanidades en general, son «blindados» tanto los nombres de los evaluadores como los de los autores.

<sup>48</sup> *Water Resources Research Unit of the University of North Carolina, Report Guidelines for Authors*, Editorial Checklist, disponible en <http://www.ncsu.edu/wrri/reports/guidelines/edchecklist.html>.

<sup>49</sup> En las revistas de filosofía, y hasta donde sé en las revistas de las humanidades en general, el tiempo permitido es mucho mayor.

<sup>50</sup> Vid., por ejemplo, LOCK y SMITH, 1990: 1342 (reportando que los resultados del estudio muestran que los evaluadores pasan menos de dos horas revisando un manuscrito); y YANKAUER, 1990: 1339 (reportado que en 12 números de la *American Journal of Public Health*, el promedio de revisión fue de 2,4 horas por evaluador por artículo —sumando en total 3,360 horas de tiempo no compensado—).

<sup>51</sup> Vid. GARDNER y BOND, 1990: 1355 (citando datos estadísticos de un estudio sobre la corrección de artículos sometidos a *The British Medical Journal*, se encontró que sólo el 11 por 100 de los artículos sometidos tenían datos estadísticos correctos y sólo el 84 por 100 de los artículos publicados eran correctos); WELLER, 1990 (reportando que muchas revistas no hacen ninguna revisión independiente de los cálculos estadísticos hechos por los autores); también vid. BRYANT *et al.*, 2006 (concluyendo que el 42 por 100 de los estudios clínicos en revistas muy bien posicionadas en el área ortopédica estaban sesgados por la inclusión de múltiples observaciones de diferentes extremidades de individuos aislados); GARCÍA BERTHOU y ALCARAZ, 2004 (encontrando que el «11,6 por 100 [21 de 181] y el 11,1 por 100 [7 de 63] de los resultados estadísticos publicados durante el 2001 en *Nature* y en *BMJ*, respectivamente, eran incongruentes» y apuntando que «[a] menos uno de tales errores aparecieron en el 38 por 100 [12 de 32] y el 25 por 100 [3 de 12] de los artículos de *Nature* y *BMJ*, respectivamente, indicando que éstos eran generalizados y no concentrados en unos pocos artículos»); NEVILLE *et al.*, 2006 (reportando que de los 364 estudios publicados de enero hasta diciembre de 2003 en los *Archives of Dermatology* y el *Journal of the American Academy of Dermatology*, «59 [38,1 por 100] de 155 [que hacían análisis estadísticos] tenían errores u omisiones en los métodos estadísticos o en la presentación de los resultados»); VAIL y GARDENER, 2003 (reportando que de los 39 estudios analizados, «[s]eis estudios eran intrínsecamente defectuosos en su diseño» y «[s]ólo 5 estudios reportaban tasas de nacidos vivos suficientes para permitir meta-análisis validos»).

<sup>52</sup> Informe presentado como *amicus brief* de D. CHUBIN *et al.* (*supra* nota 7), 11-19.

Tampoco todos los artículos que se publican en revistas con evaluación por pares son revisados por pares; algunos son invitados y algunos aparecen dado cierto privilegio editorial; y a algunos autores se les ha pedido —como alguna vez me pidieron a mí misma— que nombren a sus propios *réferis*.

En tanto que la escala de la operación crece, con más y más trabajos propuestos en más y más revistas, es muy probable que disminuya la calidad de los evaluadores y el tiempo y atención que puedan dedicarle a su trabajo. En tanto que aumenten las presiones en la carrera de los científicos, crece la tentación de los *réferis* para recomendar la aceptación de trabajos que consideran pueden promover su propia carrera, recomendar el rechazo de trabajos que perciban como una amenaza profesional o plagiar ideas de trabajos que se les ha pedido revisar<sup>53</sup>. Y en tanto incrementan las presiones sobre las revistas y su personal, la aspiración de tener prestigio y dinero causa otras distorsiones: algunas revistas suspenden el proceso de evaluación por pares cuando publican congresos que han sido patrocinados por compañías farmacéuticas (por lo cual la revista puede cobrar a la compañía una significativa factura); algunas recogen grandes sumas por la venta de un enorme número de reimpressiones de las compañías cuyos estudios han sido reportados en un artículo<sup>54</sup>; algunas presionan a los autores para que citen otros artículos publicados en la misma revista y así incrementar su «factor de impacto» y aumentar los pedidos de las bibliotecas, etcétera<sup>55</sup>.

Los propios editores han empezado a expresar su preocupación. R. SMITH, editor de *The Lancet*, escribe que la evaluación por pares es «cara, lenta, propensa a los sesgos, susceptible de abuso, quizás conservadora e incapaz de detectar un fraude»<sup>56</sup>. D. RENNIE, editor asociado de la *Journal of the American Medical Association*, es incluso más franco: «[p]arece que no hay un análisis tan fragmentado, una hipótesis tan trivial, una cita bibliográfica tan parcial o tan egoísta, ningún diseño tan deforme, ninguna metodología tan fallida, ningún argumento tan circular, ninguna conclusión tan trivial o tan poco justificada y ninguna gramática o sintaxis tan ofensiva para un papel que terminar impreso»<sup>57</sup>.

\* \* \*

Según un estudio publicado en *JAMA* en el 2004, un sondeo en 122 artículos encontró que entre sus resultados publicados, el 50 por 100 de aquellos que reportaban cierta eficacia y el 65 por 100 de aquellos que reportaban daños estaban incompletos<sup>58</sup>.

---

<sup>53</sup> En este contexto, vale la pena notar que el título de un mercedamente bien conocido artículo que critica el sistema de evaluación por pares para la financiación de proyectos: «Malice's Wonderland: Research Funding and Peer Review» (en castellano, «Malicia en el país de las maravillas: financiación de la investigación y evaluación por pares») de OSMOND (1983) [*N. de la T.*: vale la pena resaltar que el título de este artículo citado juega con el nombre del famoso libro «Alicia en el país de las maravillas» de Lewis CARROLL].

<sup>54</sup> «Dos editores reportaron que sus revistas cobraron entre \$400 y \$2,500 por página para publicar congresos, otros reportaron el cobro de una tarifa fija de \$100.000. Las revistas cobran un promedio de \$15 por reimpresión, y la reimpresión por pedido para congresos [un promedio de] 25,000» (BERO *et al.*, 1992: 1135-1140).

<sup>55</sup> BEGLEY, 2006: B1, B8.

<sup>56</sup> SMITH, 1997: 759.

<sup>57</sup> RENNIE, 1986: 2391.

<sup>58</sup> AN-WEN CHAN *et al.*, 2004: 2457.

Según un estudio comunicado por *Nature* en el 2005, más del 10 por 100 de 3.247 científicos encuestados admiten haber ocultado detalles sobre la metodología o los resultados de los trabajos; más del 15 por 100 admitió excluir las observaciones o ciertos datos; y más del 27 por 100 admitió mantener registros inadecuados de proyectos de investigación<sup>59</sup>. Según una investigación publicada en *JAMA* el mismo año, de 45 estudios ampliamente citados, publicados en prestigiosas revistas y que aducían a tratamientos médicos efectivos, 15 fueron más tarde contradichos en todo o en parte por otros estudios<sup>60</sup>. Y según un artículo publicado en la revista *Scientist* presentado por la Universidad Rockefeller en la primavera del 2006, durante los cuatro años en los que la *Journal of Cell Biology* estuvo examinando cada una de las imágenes de cada uno de los trabajos aceptados, buscando alteraciones hechas con Adobe Photoshop, 14 de 1.400 artículos fueron rechazados después de detectarse una alteración fraudulenta de la imagen<sup>61</sup>.

Aunado a esto, otros artículos sugieren que incluso después de que se ha descubierto una grave falta de ética científica o un fraude categórico, el proceso de limpieza de la literatura científica, a fin de que tales trabajos sean retractados y las referencias a éstos realizadas por otros inocentes sean corregidas es, en el mejor de los casos, irregular y desigual<sup>62</sup>. Por ejemplo, un año después de que la *Office of Research Integrity* informara a diez revistas en las que el Dr. E. POEHLMAN había publicado como co-autor que tales trabajos eran fraudulentos, sólo ocho de éstas los retractaron; e incluso después de que los *Annals of Internal Medicine* hubiesen retractado uno de dichos trabajos, otros autores continuaron citándolo<sup>63</sup>.

\* \* \*

De hecho, hay tantos reportes recientes sobre las fallas del sistema de evaluación por pares que lo difícil es seleccionar los más ilustrativos. ¿Acaso debería ser el bien conocido caso del Dr. H. WOO SUK, el investigador coreano cuyo supuesto impresionante trabajo sobre clonación, publicado en *Science* y *Nature*, resultó estar fundado en datos inventados?<sup>64</sup> ¿O debería ser aquel extraordinario artículo del *Journal of Reproductive Medicine* que afirmaba haber mostrado que una oración de intercesión por forasteros que pertenecían a otra religión en otro país doblaba la tasa de éxito de fertilizaciones in vitro logradas —cuyo supuesto autor principal luego admitió que conoció los detalles del estudio después de 6 a 12 meses de haberlo terminado y otro de los autores, un abogado sin ninguna formación científica, posteriormente se declaró culpable de los cargos (no relacionados) de fraude empre-

<sup>59</sup> MARTINSON *et al.*, 2005: 737.

<sup>60</sup> IONNADIS, 2005: 220.

<sup>61</sup> GRAVITZ, 2006, 10-12.

<sup>62</sup> FRIEDMAN, 1990: 1417; PFEIFFER y SNODGRASS, 1990: 1420 (Resumen).

<sup>63</sup> COUZIN y UNGER, 2006: 39; SOX y RENNIE, 2006: 609 [apuntando que en 1898, a efectos de evaluar las acusaciones de un fraude científico, el *Congress* creó la Oficina de la Integridad Científica (*Office of Scientific Integrity*), más tarde renombrada como Oficina de la Integridad en la Investigación (*Office of Research Integrity*)]. El artículo que fue retractado era de E. T. POEHLMAN *et al.*, «Changes in Energy Balance and Body Composition at Menopause: A Controlled Longitudinal Study», *Annals of Internal Medicine*, 123 (1995): 673-676. SOX y RENNIE, 2006: 609.

<sup>64</sup> WADE y SANG-HUN, 2006: A1 (citando al Dr. B. LEWIN, editor jefe de la revista *Cell*, comentando que la revista *Science* debería ser más cuidadosa y, sin duda, no deberían publicar un artículo con «varias fotos idénticas»).

sarial—? <sup>65</sup>. ¿O quizá los artículos de J. SUDBØ en *The Lancet* y en el *New England Journal of Medicine*, afirmando haber mostrado que los fármacos antiinflamatorios no esteroideos reducían el riesgo del cáncer bucal, todo lo cual resultó estar fundado en datos inventados? <sup>66</sup>. ¿O debe ser algo menos fuerte, como el artículo del *New England Journal of Medicine*, citado en el proceso seguido contra Metabolife, en el cual la información en una tabla de 11 pacientes que enlistaba los efectos adversos y las condiciones preexistentes es contradicha por el texto en la misma página? <sup>67</sup>.

No, ninguno de éstos, sino la extraordinaria saga del reporte sobre el estudio clínico a gran escala del fármaco Vioxx, hecho por Merck: el estudio VIGOR es claramente el ejemplo más instructivo de lo que puede ir mal en este proceso <sup>68</sup>. Después de la aprobación de la FDA, el reporte del estudio —concluyendo que Vioxx conllevaba un riesgo más bajo de provocar efectos gastrointestinales adversos que otros analgésicos y que para la mayoría de los pacientes el riesgo de efectos cardiovasculares adversos no era significativo— fue enviado al *New England Journal of Medicine*, donde apareció en noviembre del 2000 <sup>69</sup>. Sin embargo, en el 2002 Merck fue obligada a añadir en cada caja del producto una advertencia sobre sus riesgos cardiovasculares. Y en septiembre del 2004 —después de que el *data safety monitoring board* detuviera otro estudio clínico a gran escala, el estudio APPOVe (diseñado para demostrar que el Vioxx reducía el riesgo de pólipos en el colon), cuando se supo que los pacientes que recibieron 25 mg de Vioxx durante más de 18 meses tuvieron una incidencia cuatro veces mayor de sufrir graves episodios tromboembólicos— Merck retiró del mercado al fármaco <sup>70</sup>.

En diciembre del 2005, en medio de la creciente tormenta de litigios iniciados por pacientes afirmando haber sufrido daños por el fármaco, la *New England Journal of Medicine* (en adelante, *NEJM*) hizo pública una «*Expression of Concern*» («declaración de preocupación») reconociendo que tres infartos sufridos por pacientes que habían tomado Vioxx fueron omitidos en el reporte del estudio VIGOR publicado en el 2000 <sup>71</sup>. Estos adversos eventos habían sido incluidos desde febrero del 2001 en la

<sup>65</sup> CAREY, 2006: A15; FLAMM, 2004: 2-38. El artículo implicado era de K. W. CHA *et al.*, «Does Prayer Influence the Success of In Vitro Fertilization-Embryo Transfer? Report of a Masked, Randomized Trial», *Journal of Reproductive Medicine*, 2001, 46, 781-787, 782. FLAMM, 2004: 27-28.

<sup>66</sup> SUDBØ *et al.*, 2005: 1359-1366. La base de datos de 908 participantes en el estudio, supuestamente, fue simplemente inventada; ¡250 de las personas ficticiamente involucradas tenían la misma fecha de nacimiento! FORBES, «Muchos investigadores rompieron las reglas: estudio», en <http://www.forbes.com/forbeslife/health/feeds/bscout/2006/04/13/bscout532110.html> (consultado el 13 de abril de 2006). *Vid.* también HORTON, 2006: 196 (expresando preocupación sobre la admisión verbal que SUDBØ hizo sobre la fabricación de datos para el estudio previamente publicado en el *Lancet* y reconociendo su posible conducta indebida en dos de sus artículos de investigación publicados en el *New England Journal of Medicine*); HORTON, 2006: 382 (retirando el artículo de SUDBØ, dada la confirmación de que los datos fueron fabricados).

<sup>67</sup> HALLER y BENOWITZ, 2000: 1836. La tabla 4 de la página 1836 registra al paciente número siete como no teniendo afecciones pre-existentes o riesgos asociados, mientras en el texto en la misma página indica que la autopsia del paciente «mostró cardiomegalia moderada con dilatación de cuatro cámaras cardíacas y enfermedad de la arteria coronaria, con un estrechamiento del 50 al 75% de cuatro vasos».

<sup>68</sup> *Senate Committee on Finance, FDA, Merck, and Vioxx: Putting Patient Safety First?* 108th Cong. (Nov. 18, 2004) (testimonio de la Dra. S. KWEDER, directora de la oficina de drogas nuevas en el Centro para la Evaluación e Investigación de Drogas). Disponible en <http://www.finance.senate.gov/hearings/bearing/?id=48b33994-9de9-df09-dfb8-bb3599a92fbb>.

<sup>69</sup> BOMBADIER *et al.*, 2000: 1520.

<sup>70</sup> MAXWELL y WEBB, 2005: 449.

<sup>71</sup> ARMSTRONG, 2006: A1, A10.

información que aparecía en la página web de la FDA; y dos de los tres autores sabían de éstos mucho antes de la publicación del artículo<sup>72</sup>. Su inclusión aumentó la tasa de los infartos entre aquellos que tomaban Vioxx, del 0,4 por 100 afirmado en el artículo al 0,5 por 100 (comparado con el 0,1 por 100 entre pacientes que toman naproxeno) y además contradiciendo también la afirmación de que sólo aquellos que ya tenían riesgo mostraban un incremento de infarto después de tomar Vioxx<sup>73</sup>. Merck afirmó que los infartos adicionales ocurrieron después de la fecha de corte del estudio; pero el jefe de redacción de la revista, Dr. J. DRAZEN, le dijo a los periodistas que el diseño del estudio, que él había seguido para hacer el seguimiento de los efectos gastrointestinales después de que se detuvo el seguimiento de los efectos cardiovasculares, inducía a errores<sup>74</sup>.

Pero el problema aquí no fue sólo con los autores; tampoco fue sólo que los evaluadores no hayan tenido disponibles los datos brutos o que hubiesen errado al no observar algunos aspectos extraños en el diseño del estudio. Ahora sabemos que en junio del 2001 los editores del *NEJM* habían recibido una carta de la farmacéutica J. HRACHOVEC solicitando la corrección del artículo en vista de la información publicada en la página web de la FDA, pero se negó su publicación porque la revista «no podía estar controlando toda información publicada»<sup>75</sup>; también sabemos que el director ejecutivo, Dr. G. CURFMAN, en noviembre del 2005, presentado por las partes durante una deposición en un juicio federal en Texas, reconoció que ni los evaluadores ni los editores habían cuestionado la teoría de Merck sobre una tasa más alta de eventos cardiovasculares entre los pacientes Vioxx atribuible al efecto cardioprotector del naproxeno, a pesar de que un oficial de la FDA había observado que esto «no está fundamentado por ningún [...] estudio clínico controlado»<sup>76</sup>; que la revista había vendido 929.000 copias de reimpresiones del artículo, la mayoría de ellas a Merck, con unos ingresos estimados entre 697.000 y 836.000 dólares; y también sabemos que la «declaración de preocupación» sobre el estudio había sido publicada inmediatamente después de un urgente comunicado de último minuto de E. CAFASSO, especialista en relaciones públicas, diciendo que dado el testimonio que se presentaría al día siguiente en el caso Vioxx, donde el Dr. CURFMAN había depuesto, era indispensable para la revista publicar algo de forma inmediata a efectos de «desplazar a los medios de comunicación del *NEJM* hacia los autores, Merck y los abogados de los demandantes»<sup>77</sup>. Como observó R. SMITH, exdirector de la *British Medical Journal*, la conducta de la *New England Journal* en la disputa sobre el estudio VIGOR «generó dudas sobre la integridad de la revista»; «la revista le falló a sus lectores [y] dañó su reputación»<sup>78</sup>.

<sup>72</sup> *Ibid.*

<sup>73</sup> *Ibid.*

<sup>74</sup> *Ibid.*

<sup>75</sup> *Ibid.* (citando al Dr. J. DRAZEN, un editor jefe del *New England Journal of Medicine* durante una declaración en radio en agosto del 2001).

<sup>76</sup> *Ibid.*

<sup>77</sup> *Ibid.* Una expresión de preocupación fue publicada en línea el 8 de diciembre de 2005, el día en que el jurado inició su deliberación en el tercer juicio de Vioxx (HENDERSON, 2005). De acuerdo con HENDERSON, en diciembre de 2005, el Dr. CURFMAN dijo que la *NEJM* había «aprendido de la nueva información [es decir, los tres ataques al corazón omitidos] hacía dos semanas [antes]»; de acuerdo con ARMSTRONG, sin embargo, la revista sabía de ello desde la carta del Sr. HRACHOVEC en junio del 2001.

<sup>78</sup> Según el *Medical News Today*, «el *New England Journal of Medicine* está dañado por su conducta en el caso VIOXX, dijo el editor jefe del *British Medical Journal*». Vid. en [www.medicalnewstoday.com/medicalnews.php?newsid=46831](http://www.medicalnewstoday.com/medicalnews.php?newsid=46831) (revisado el 9 de julio del 2006).



Y justo cuando se pensaba que no podía ser peor, en julio del 2006 la *NEJM* publicó una corrección al reporte sobre el estudio APPROVe que anteriormente había publicado: los principales resultados afirmados en el artículo no habían sido alcanzados por el método estadístico que los autores afirmaban haber usado; es más, usando el método que los autores afirmaban haber usado, pero que de hecho no usaron, los resultados socavaban que los incrementos de riesgos cardiovasculares empiezan después de 18 meses de tomar el fármaco, como se afirmaba en el artículo<sup>79</sup>.

No mucho tiempo antes, L. ALTMAN había escrito en el *New York Times* que «recientes revelaciones de estudios fraudulentos o con errores en las revistas médicas y científicas, habían puesto en cuestión como nunca antes [al sistema de evaluación por pares] [...]»<sup>80</sup>; es difícil no estar de acuerdo.

\* \* \*

Por obvias razones estos estudios son frecuentemente más difíciles de rastrear y, también por obvias razones, son sólo conocidos hasta mucho tiempo después de la revelación; sin embargo, es muy claro que hay muchos ejemplos de trabajos importantes e innovadores que han sido rechazados por los evaluadores. LOCK cuenta la historia del reporte de E. JENNER sobre su vacuna contra la viruela, que en 1796 fue rechazado por la revista *Transactions* de la *Royal Society*, después de que Sir J. BANKS lo revisara y reportara que él «no confiaba» en su conclusión<sup>81</sup>. Ch. MCCUTCHEON, lamentando la forma en que «los evaluadores rechazan manuscritos tanto buenos como pobres», enlista a «Frederick Lanchester con su teoría de la circulación sobre cómo las alas se sostienen, planteada en 1894, las estadísticas del fotón de Chandra Bose, la teoría de la emisión beta de Enrico Fermi en 1933, el descubrimiento de Herman Almqvist de la vitamina K<sub>2</sub> en 1935, el ciclo del ácido nítrico de Hans Krebs en 1937 y el concepto de dinámica trófica en ecología propuesto por Raymond Lindeman en 1941»; todos fueron «rechazados al menos una vez»<sup>82</sup>. D. HORROBIN añadió que el artículo de KREBS «posiblemente el artículo más importante en la bioquímica moderna, [...] finalmente le llevó a recibir un Premio Nobel»<sup>83</sup>; y enlista muchos otros ejemplos, incluyendo: un artículo de gran influencia [...] en inmunología de GLICK *et al.*, sobre la identificación de los linfocitos B, que fue «rechazado por las más importantes revistas generales y especializadas y que finalmente apareció en *Poultry Science*, debido a la especie en que se realizó la investigación»<sup>84</sup>; y un artículo de la granjera neozelandesa G. REID, sugiriendo que el eczema facial en las ovejas podría ser causado por una marginal deficiencia de zinc, que fue rechazado por las revistas del área, hasta que HORRIBIN lo publicara en *Medical Hypotheses* —después de que su trabajo fue confirmado, la enfermedad fue

<sup>79</sup> WON TESORIERO, 2007: A2. El artículo original era de BRESALIER *et al.*, 2005: 1092-1102. Por irónico que parezca, el mismo número de la revista incluye un artículo pequeño de J. M. DRAZEN titulado «COX-2 Inhibitors-A Lesson in Unexpected Problems», *New England Journal of Medicine*, 352 (2005), 1131-1132. El siguiente año, BRESALIER *et al.*, publicó una corrección: «Correction», *New England Journal of Medicine*, 355 (2006), 221.

<sup>80</sup> ALTMAN, 2006: F1.

<sup>81</sup> LOCK, 1985: 2

<sup>82</sup> MCCUTCHEON, 1991: 33.

<sup>83</sup> HORROBIN, 1990: 1440 [este artículo fue citado en *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579, 593 (1993)].

<sup>84</sup> HORROBIN, 1990: 1440.

eliminada y a la Sra. REID le fuera concedida una condecoración por sus servicios a la agricultura de Nueva Zelanda—<sup>85</sup>.

A este punto ya casi no debería ser necesario decir que: «El hecho de que un trabajo haya pasado el proceso de evaluación por pares para su publicación no garantiza que no tenga defectos o incluso que no sea fraudulento; y el hecho de que el trabajo haya sido rechazado por evaluadores no garantiza que no sea un progreso importante».

#### 4. LECCIONES PARA ABOGADOS

«¡Basta ya!», usted debe estar pensando. Ciertamente, la confianza del juez KOZINSKI en que «el proceso normal de evaluación por pares [...] es uno de los distintivos de la investigación científica fiable» fue demasiado optimista<sup>86</sup>; pero, ¿el magistrado BLACKMUN no clarificó esto, hace dos décadas atrás, en su criterio en el caso *Daubert*?<sup>87</sup>.

Bueno, evidentemente el magistrado BLACKMUN atendió a los argumentos del *amici* CHUBIN *et al.*<sup>88</sup>, al reconocer que:

[L]a publicación (que es sólo un elemento de la evaluación por pares) no es condición *sine qua non* para la admisibilidad; no está relacionada necesariamente con la fiabilidad [...] y, en algunos casos, no son publicadas teorías que pese a estar bien fundadas son innovadoras [...].

Pero someter al escrutinio de la comunidad científica [...] incrementa las probabilidades de que se detecten defectos sustantivos en la metodología<sup>89</sup>.

Hubiera sido deseable plantear más explícitamente la distinción entre el sentido amplio y el sentido estricto de «evaluación por pares»; sin embargo, lo que el magistrado BLACKMUN tiene en mente parece razonablemente claro; y, además, parece cierto: trabajos científicos pobres pueden pasar el proceso de evaluación por pares y trabajos científicos buenos pueden no hacerlo, pero cuando es publicado un trabajo científico y queda disponible al escrutinio de otros científicos, incrementa la probabilidad de que a final de cuentas cualquier defecto metodológico serio será detectado. Y el consejo del magistrado BLACKMUN sobre el peso que los tribunales deberían dar a este «factor *Daubert*» —en efecto, esta es una consideración pertinente pero no decisiva— parece difícil negar:

El hecho de que un artículo sea publicado en una revista con *referato*, o que no sea publicado en ésta, será una consideración relevante, aunque no determinante, para valorar la validez científica de una específica técnica o metodología que sirve de fundamento a una inferencia experta<sup>90</sup>.

<sup>85</sup> *Ibid.*

<sup>86</sup> *Daubert v. Merrell Dow Pharms., Inc.*, 951 F.2d 1128, 1131 n.3 (9th Cir. 1991).

<sup>87</sup> *Vid. Daubert* (1993), 509 U.S. 593 (estableciendo que tomar en cuenta si una idea ha sido sujeta a la evaluación por pares es una «consideración pertinente»).

<sup>88</sup> Dada la presión para ahorrar dinero y publicar a tiempo, los editores de revistas con *referato* pueden alguna vez publicar «el equivalente científico a un tabloide de supermercado» (Informe presentado como *amici brief* de CHUBIN, *supra* nota 7, p. 12). Un compromiso entre la «absoluta certeza» en la validez de las afirmaciones científicas en los artículos y «la absoluta velocidad y la total economía» significa que «los errores se convierten en inevitables y que reportes erróneos, engañosos y fraudulentos sean a veces publicados», *ibid.*

<sup>89</sup> *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579, 594 (1993).

<sup>90</sup> *Ibid.*, 594.

«A primera vista» (aunque a segunda vista usted se encontrara asaltado por preocupaciones tanto teóricas como prácticas): el significado de «fiable» amenaza con empezar a parecer cada vez menos determinado, cada vez más ambiguo; y el consejo a los tribunales sobre usar la evaluación por pares para determinar la fiabilidad suena cada vez menos útil. Las ambigüedades golpean a uno casi inmediatamente: ¿los tribunales deben preguntar si el trabajo que sirve como base al testimonio experto ofrecido se ha publicado después de haber sobrevivido la evaluación por pares o es suficiente con que sea publicado en «una revista con *referato*»? ¿el trabajo *del perito* debe haber sido sujeto a la evaluación por pares y publicado o es suficiente con que el perito se base en un trabajo *de otros* arbitrado y publicado?, etcétera.

La intuición del magistrado BLACKMUN sobre que a largo plazo sobrevivir al escrutinio de la comunidad científica es más o menos el mejor indicador de validez científica que un lego puede tener, aunque sea un indicador falible, es totalmente correcta; pero no es de genuina ayuda práctica. Por obvias razones, las cuestiones científicas en juego en los procesos judiciales probablemente no usan la ciencia más firmemente establecida, sino cuestiones aún controvertidas; y sería poco realista imaginar que los tribunales de alguna manera pueden decidir qué afirmaciones científicas, aún controvertidas, terminarán *sobreviviendo* tal «evaluación por pares», cuando los propios científicos no pueden hacerlo.

Y en lugar de clarificar el concepto de «fiabilidad probatoria» (que fue equiparado con «validez científica» por la Corte Suprema en el caso *Daubert*)<sup>91</sup>, las observaciones del magistrado BLACKMUN contribuyeron a su confusión<sup>92</sup>. En el lenguaje ordinario «fiable» tiene una compleja maraña de usos: pero tanto si estamos describiendo objetos inanimados, como relojes o carros, personas («informantes» o «fuentes», como solemos decir), o información, bases de datos, etc., la fiabilidad ordinariamente es concebida como una cuestión gradual. Sin embargo, el criterio *Daubert* es sobre la admisibilidad de pruebas periciales, que no es una cuestión gradual; y ello entonces nos obliga a adoptar una concepción categórica.

Si las pruebas deben ser lo *suficientemente* fiables para ser admisibles, ¿cuán fiables deben ser y cómo un tribunal determinará si las pruebas alcanzan dicho estándar? (¿se debe imponer el mismo grado de fiabilidad a las pruebas científicas «blandas» que a las «duras» o a las pruebas periciales no científicas que a las científicas?) Parece tener sentido, como el magistrado BECKER argumentó en el caso *Paoli*, que «el estándar probatorio de fiabilidad [debe ser] más bajo que el estándar de suficiencia probatoria»<sup>93</sup>; si el estándar de admisibilidad fuese tan alto como el estándar de prueba, a una parte que intenta introducir su prueba pericial se le exigiría, en efecto, probar dos veces su caso —y el tribunal estaría haciendo la tarea que corresponde hacer al jurado—. Pero ahora uno empezaría a preguntarse: ¿la publicación arbitrada, después de todo, es

<sup>91</sup> *Ibid.*, 590-591.

<sup>92</sup> «Muchos factores pesan sobre la investigación y no asumimos proponer un listado o análisis definitivo. Pero algunas observaciones generales son apropiadas». *Ibid.*, 593. Bajo *Daubert* (1993), los factores clave para determinar la fiabilidad probatoria, o la validez científica, son si una teoría o técnica puede ser y ha sido sometida a prueba, si ha sido sometida a una evaluación por pares, el rango de error conocido o potencial, los estándares que regulan la operación de la técnicas, y «la aceptación general». *Ibid.*, 593-594.

<sup>93</sup> *In re Paoli R.R. Yard PCB Litig.*, 35 F.3d 717, 744 (3d Cir. 1994).

suficiente para garantizar que las pruebas ofrecidas alcanzan cierto umbral mínimo como estándar de fiabilidad? Y si no lo es, al menos es suficiente para garantizar que la metodología empleada cumple los estándares mínimos, incluso si las conclusiones planteadas no son fiables?, ¿no es eso lo que el juez KOZINSKI tenía en mente cuando en 1995 escribió que la publicación arbitrada «es un importante indicador de que se cumple *al menos con los criterios mínimos* de la buena ciencia»?<sup>94</sup>.

El criterio del magistrado BLACKMUN deja todo esto abierto<sup>95</sup>. La opinión del magistrado REHNQUIST en el tribunal del caso *Joiner*, indicando cierta incertidumbre sobre la distinción entre la metodología y las conclusiones usada en el caso *Daubert*, no ayuda<sup>96</sup>. Y la opinión del magistrado BREYER en el tribunal del caso *Kumho* —sosteniendo que la tarea de guardianes impuesta por *Daubert* se extendía tanto a las pruebas periciales no científicas como a las científicas—\*, pero que los tribunales podían usar alguno, todos o ninguno de los factores *Daubert* y/u otros factores más apropiados para la tarea en cuestión —confirma que las cuestiones más delicadas se dejan a la discrecionalidad judicial—<sup>97</sup>.

Por supuesto, no es sorprendente que la Corte en *Daubert* no encontrara una fórmula precisa para decidir sobre la fiabilidad probatoria; e incluso si tal cosa fuera posible, probablemente no sería la precisión deseable sino el tipo de «exactitud engañosa» que una vez O. WENDELL HOLMES condenó como «una fuente de falacia en todo el derecho»<sup>98</sup>. Y, especialmente porque «la evaluación por pares y la publicación» es sólo uno de los factores en la flexible lista *Daubert*, quizá tampoco es sorprendente no encontrar una clara correlación entre las decisiones que admiten o excluyen pruebas periciales como resultado de si han o no satisfecho este factor. En cambio:

i) Algunos tribunales (citando el reconocimiento del magistrado BLACKMUN de que la publicación arbitrada no es condición *sine qua non* para la admisibilidad) han admitido pruebas periciales que no tienen fundamento en trabajos arbitrados y publicados<sup>99</sup>.

<sup>94</sup> *Daubert v. Merrell Dow Pharms., Inc.*, 43 F.3d 1311, 1318 (9th Cir. 1995) (cursiva añadida).

<sup>95</sup> *Supra* notas 91-94 y el texto que les acompaña.

<sup>96</sup> *Gen. Elec. Co. v. Joiner*, 522 U.S. 136, 146 (1997) («las conclusiones y la metodología no son totalmente distintos uno del otro»).

\* *N. de la T.*: la «tarea de guardianes impuesta por *Daubert*» hace referencia a que la Corte Suprema estadounidense en éste caso enfatizó que los jueces tienen que valorar (además de su pertinencia) la fiabilidad de las pruebas científicas *para su admisión*. Los jueces entonces servirían de «guardianes» (en inglés se usó la palabra «gatekeepers») de la calidad de las pruebas científicas que entran a un proceso judicial y que serán conocidas por el jurado. La Corte además sugirió un listado de criterios para valorar dicha fiabilidad, mismos que han sido citados previamente por la autora en la nota 92.

<sup>97</sup> *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 141 (1999). La Corte sostuvo que:

La determinación general del caso *Daubert* —estableciendo la obligación a todos los jueces de actuar como guardianes— aplica no sólo al testimonio basado en conocimiento «científico» sino también al testimonio basado en conocimiento «técnico» y «especializado» [...] [U]n juez puede considerar uno o más de los factores específicos mencionados en *Daubert* siempre que hacerlo sea de ayuda [...]. Pero [...] la lista de los factores específicos en *Daubert* no necesariamente ni exclusivamente es de aplicación para todos los expertos en cada caso.

<sup>98</sup> *Gompers v. United States*, 233 U.S. 604, 610 (1914).

<sup>99</sup> *Vid.*, por ejemplo, *Ruiz Troche v. Pepsi Cola of P.R. Bottling Co.*, 161 F.3d 77, 84 (1st Cir. 1998) (revocando la exclusión del testimonio del Dr. O'DONNELL, decretada por el tribunal de distrito, respecto los efectos de la cocaína en el comportamiento de un conductor, sobre la base de que, pese a que las fuentes secundarias citadas por él no habían sido sometidas a la evaluación por pares o publicadas, otros estudios sometidos a la

ii) Algunos tribunales (citando el reconocimiento del magistrado BLACKMUN respecto a que la publicación arbitrada no está necesariamente relacionada con la fiabilidad) han excluido pruebas periciales basadas en trabajos arbitrados y publicados<sup>100</sup>.

iii) Algunos tribunales (citando el reconocimiento del magistrado BLACKMUN de que la publicación arbitrada es una consideración pertinente) en parte han admitido una prueba porque se basaba en un trabajo arbitrado y publicado<sup>101</sup>.

iv) Algunos tribunales (también citando el reconocimiento del magistrado BLACKMUN de que la publicación arbitrada es un factor pertinente) en parte han excluido pruebas porque no tenían tal base<sup>102</sup>.

Además no es muy sorprendente, dados los cambios del sentido amplio al sentido estricto de «evaluación por pares» del magistrado BLACKMUN, que algunos tribunales hayan interpretado «la evaluación por pares» para incluir algún tipo de exposición a otras personas del área, distinta al proceso de evaluación por pares para la publicación<sup>103</sup>. Tampoco es nada sorprendente que «la evaluación por pares y la publicación»

---

evaluación por pares y publicados apuntaban lo mismo); *Kannankeril v. Terminix Int'l, Inc.*, 128 F.3d 802, 809 (3d Cir. 1997) [anulando y reenviando la decisión al tribunal inferior que había: i) excluido el testimonio del Dr. GERSON argumentando que «aunque el Dr. Gerson nunca había escrito en el tema, su opinión estaba fundamentada en amplios conocimientos científicos aceptados sobre la naturaleza dañina de los fosfatos orgánicos»; y ii) observando que en el caso *McCulloch v. H.B. Fuller Co.*, 61 F.3d 1038, 1042 (2d Cir. 1995) se sostuvo que la evaluación por pares, la publicación y la aceptación general serían considerados en la valoración de la prueba y no para su admisibilidad]; *Metabolife Int'l, Inc. v. Wornick*, 264 F.3d 832, 843 (9th Cir. 2001) [citando *Daubert* (1995), 1317, respecto la afirmación de que «cuando la investigación ha iniciado antes del litigio, podría ser fiable aunque no estuviese sometida a la revisión por pares»]; *United States v. Hankey*, 203 F.3d 1160, 1168 (9th Cir. 2000) (sosteniendo que el tribunal de distrito no abuso de su discrecionalidad al admitir el testimonio de un policía experto en códigos de las pandillas y citando *Kumbo Tire*, 526 U.S., 152, al decir que un tribunal debe tener libertad no sólo para decidir si admite o no un testimonio experto sino también para decidir «cómo valorar la fiabilidad del experto»).

<sup>100</sup> *Vid.*, por ejemplo, *Allison v. McGhan Med. Corp.*, 184 F.3d 1300, 1313, 1316, 1319, n.24 (11th Cir. 1999) (manteniendo la exclusión de los testimonios expertos resuelta por el tribunal inferior sobre la función del silicón en los implantes mamarios que causaron los daños de las demandantes, fundándose en parte en el hecho de que un estudio que había sido sometido a la evaluación por pares y publicado «no significa que constituyese una base adecuada» para la inferencia experta, que «el escrutinio realizado por los pares de un experto no aseguraba la admisibilidad de la prueba», y que el hecho de que un perito hubiese publicado muchos artículos en revistas con *referato* «no establecía la validez científica de su premisa»); *United States v. Cordoba*, 194 F.3d 1053, 1059 (9th Cir. 1999) (manteniendo la exclusión de la prueba poligráfica decidida por el tribunal de distrito, aun cuando cientos de artículos sobre el polígrafo habían sido publicados, incluso en muchas revistas con *referato*).

<sup>101</sup> *Vid.*, por ejemplo, *In re Silicone Gel Breast Implants Prods. Liab. Litig. (California)*, 318 F. Supp. 2d 879, 896 (C.D. Cal. 2004) (resolviendo la admisibilidad del análisis y crítica realizados por el Dr. NEUGEBAUER sobre la existencia de pruebas epidemiológicas, en parte porque «[l]as claves estadísticas de la epidemiología [...] han sido sujetas a la evaluación por pares y su publicación»).

<sup>102</sup> *Vid.*, por ejemplo, *Berry v. City of Detroit*, 25 F.3d 1342, 1350-1351 (6th Cir. 1994) (excluyendo el testimonio de Leonard Postill sobre el entrenamiento policial presentado por los demandantes, fundado en parte en que «[c]iertamente no hay ningún testimonio como una revisión por pares de la teoría de Postill»); *Nat'l Bank of Commerce v. Associated Milk Producers, Inc.*, 191 F.3d 858, 864-865 (8th Cir. 1999) [confirmando la exclusión resuelta por el tribunal inferior del testimonio experto que relacionaba la aflatoxina M-1 (AFM) y el cáncer del demandante, con el argumento de que «[n]o hay estudios científicos o literatura médica que muestre ninguna correlación entre el ADM y el cáncer de laringe»]; *Nelson v. Tenn. Gas Pipeline Co.*, No. 95-1112, 1998 WL 1297690, \*8-9, \*13 (W. D. Tenn. Aug. 31, 1998) (excluyendo el testimonio de los expertos del Sr. NELSON, Drs. KILBURN y HIRSCH, que afirmaban que los daños del demandante habían sido causados por su exposición al PCB de un gasoducto, con el argumento de que su trabajo no había sido publicado ni sometido a la evaluación por pares).

<sup>103</sup> *Vid.*, por ejemplo, *United States v. Bonds*, 12 F.3d 540, 559-560 & n. 16, 568 (6th Cir. 1993) (confirmando la decisión del tribunal inferior sobre la admisión del testimonio de un experto del FBI sobre ADN,



haya encontrado un espacio en los tribunales estatales estadounidenses que no han adoptado el criterio *Daubert*<sup>104</sup> e incluso en algunos casos que suponen cuestiones muy diferentes a la admisibilidad de la prueba pericial<sup>105</sup>.

Pero es decepcionante encontrar que el análisis que los tribunales han hecho de «la evaluación por pares y la publicación» parece haber sido, en su mayoría, de poca profundidad; porque nuestra investigación de las virtudes y los vicios del sistema de evaluación por pares ha sugerido toda una serie de preguntas que podrían arrojar luz sobre la importancia de que una prueba pericial ofrecida en un caso se base, o no, en un trabajo publicado en una revista con *referato*. ¿Epistemológicamente cuán respetable es el área en cuestión<sup>106</sup>, y actualmente hay desacuerdos metodológicos serios en ésta?, ¿es ésta una revista muy bien considerada en el área o es de segundo o tercer nivel —o es el último recurso de aquellos desesperados por publicar—?, ¿el trabajo publicado en una «revista con *referato*» fue realmente evaluado o fue publicado por privilegio editorial o invitado? Si fue arbitrado, ¿los revisores fueron sugeridos por el autor (o autores)? Si fue invitado, ¿fue por la buena reputación del autor o por su relación personal con el editor?, ¿el autor tiene relación con la revista, por ejemplo, es parte del comité editorial?, ¿la revista en la que se publicó el trabajo recibe apoyo, directa o indirectamente, de alguna de las partes de este proceso concreto o está estrechamente relacionado con otros litigios del mismo tipo?, ¿el trabajo fue rechazado por otras revistas antes de ser aceptado por ésta y, de ser así, por cuántas, por cuáles y con qué fundamentos? Si la prueba pericial se funda en un trabajo que no ha sido publicado, ¿es porque éste es muy reciente o porque, aun cuando no es tan reciente, nunca fue sometido a publicación o porque fue sometido pero rechazado?, ¿ha habido subsecuentes declaraciones de preocupación o retractaciones<sup>107</sup> u otros trabajos han criticado este otro?

---

aun cuando «muchos de los artículos introducidos como [...] publicados no lo estaban en una revista con *referato* en el sentido estricto del término», puesto que «todos los artículos sometidos a los procedimientos del FBI son expuestos a la comunidad científica»); *United States v. Havvard*, 117 F. Supp. 2d 848, 854 (S.D. Ind. 2000) (admitiendo el testimonio del FBI sobre la identificación mediante huellas dactilares, argumentando y concluyendo que satisfacía *Daubert*; en particular los métodos de un examinador de huellas dactilares son sujetos a la revisión por pares porque «cualquier otro examinador cualificado puede comparar la información objetiva en la que estaban basados las inferencias expertas y podría ofrecer inferencias expertas distintas si estuviesen justificadas»). *Havvard*, creo, extiende el significado de «revisión por pares» mucho más allá de sus límites razonables.

<sup>104</sup> *Vid.*, por ejemplo, *Berry v. CSX Transp., Inc.*, 709 So. 2d 552, 569-570 (Fla. Dist. Ct. App. 1998) (argumentando que, incluso bajo *Frye*, «[a]unque la existencia de numerosos estudios sometidos a la evaluación por pares y publicados no garanten que tales estudios no tienen errores, tal publicación [...] mitiga la necesidad de un escrutinio judicial... en la etapa de admisibilidad»).

<sup>105</sup> *Vid.*, por ejemplo, *Kitzmiller v. Dover Area Sch. Dist.*, 400 F. Supp. 2d 707, 743-745 (M. D. Pa. 2005) (decidiendo que, mientras el testimonio experto presentado por el demandante esté basado en literatura sometida a la evaluación por pares, los testimonios expertos del demandado no estén basados en material que ha sido sujeto a la evaluación por pares, lo que es «espléndidamente importante» en el proceso científico, ayudando a asegurar «que los artículos sobre alguna investigación son científicamente precisos [...], satisfacen los estándares del método científico y son relevantes para otros científicos en el área»).

<sup>106</sup> *Vid.* *Kumbo Tire* (*supra* nota 75), 151 (1999) («[t]ampoco [...] la satisfacción del factor de la aceptación general enunciado en *Daubert* ayuda a mostrar que el testimonio experto es fiable cuando la disciplina en sí misma carece de fiabilidad»).

<sup>107</sup> Para encontrar retracciones en las revistas médicas, consúltase PubMed (disponible en <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed&itool=toolbar>) o se puede buscar «retracted publication» en la base de datos (*Medical Subject Heading*), luego dar click en *Links* y seleccionar PubMed.



Estas no son preguntas fáciles de responder y no es de destacarse que los tribunales no las planteen rutinariamente. Pero cuando algunas de éstas *fueron* investigadas por un tribunal —como sucedió en otro caso relacionado con el Bendectin, menos conocido que *Daubert*— los resultados son, por decir lo menos, instructivos; y algo preocupantes.

## 5. ¿VOLVEMOS AL INICIO? LA EVALUACIÓN POR PARES EN LA LITERATURA SOBRE EL BENDECTIN

El caso *Blum v. Merrell Dow Pharmaceuticals*<sup>108</sup> fue un caso absurdamente largo en los tribunales de Pensilvania que inició muchos años antes de *Daubert*, en 1982, pero que se terminó finalmente hasta el 2000<sup>109</sup>.

En el caso *Blum* como en *Daubert*, los abogados de Merrell Dow argumentaron que las pruebas periciales de los demandantes deberían ser excluidas porque no eran aceptadas por la generalidad de la comunidad científica relevante<sup>110</sup>. Sin embargo, los abogados de Blum argumentaron que *los peritos de Merrell Dow* deberían ser excluidos porque el supuesto «consenso científico» sobre la cuestión era totalmente artificial; que, de hecho, había sido creado por los fabricantes demandados al otorgar soporte a la investigación que era favorable a su fármaco y —el punto principal dados nuestros objetivos— el soporte de Merrell Dow a revistas supuestamente arbitradas que publicarían resultados útiles a la compañía a efectos de defenderse en los litigios interpuestos por el Bendectin<sup>111</sup>.

La sentencia del juez BERNSTEIN en el segundo juicio en 1996 incluye un sumario devastador del testimonio de los peritos de Merrell Dow<sup>112</sup>. Un perito de la demandada, el Dr. BRACKEN, reconoció no sólo que artículos que son «menos que buenos» pueden pasar la evaluación por pares sino que su propio estudio publicado sobre el Bendectin y los defectos congénitos fue menos que bueno<sup>113</sup>. Otro perito, Dr. KLEBANOFF, testificó que el Bendectin no causa defectos congénitos pero después admite que su artículo mostró una correlación estadísticamente significativa con cataratas congénitas, pulmones subdesarrollados y microcefalias<sup>114</sup>. El experto del demanda-

<sup>108</sup> *Blum v. Merrell Dow Pharm., Inc.*, 1 Pa. D. & C.4th 634 (Phila. Cnty. Ct. 1988) (en adelante, *Blum* 1988), *rev'd*, 560 A.2d 212 (Pa. Super. Ct. 1989) (en adelante, *Blum* 1989), *aff'd*, 626 A.2d 537 (Pa. 1993) (en adelante, *Blum* 1993), *remanded to* 33 Phila. Cnty. Rep. 193 (1996) (en adelante, *Blum* 1996), *rev'd*, 705 A.2d 1314 (Pa. Super. 1997) (en adelante, *Blum* 1997), *aff'd*, 764 A.2d 1 (Pa. 2000) (en adelante, *Blum* 2000).

<sup>109</sup> *Vid.* HAACK (2008) para más detalles de este caso fascinante, aunque perturbador. Pennsylvania, por cierto, no ha adoptado *Daubert* sino que sigue siendo un Estado en donde es aplicable el criterio *Frye*. Una revisión de la regla probatoria 702 entró en vigor en marzo del 2013; pero el texto fundamental siguió siendo si «la metodología empleada por el experto es aceptada por la generalidad en el área».

<sup>110</sup> *Blum* 2000, *supra* nota 108.

<sup>111</sup> *Blum* 1997, *supra* nota 108, 1314, 1319.

<sup>112</sup> *Vid.* *Blum* 2000, *supra* nota 108, 10 (voto particular en contra de J. CASTILLE).

<sup>113</sup> *Blum* 1996, *supra* nota 108, 207-208.

<sup>114</sup> *Ibid.*, 208. El estudio del Dr. BRACKEN estaba fundado en entrevistas con 1.427 madres, de las cuales 122 habían tomado Bendectin durante su embarazo, y éste mostraba un incremento estadístico significativo del riesgo de daños congénitos cuando una madre había tomado Bendectin y fumado.

<sup>115</sup> *Ibid.*, 209. En la examinación cruzada, el Dr. KLEBANOFF testificó que la asociación positiva entre el Bendectin y los pies equinovaros, aunque no estadísticamente significativo, satisfacía el estándar que él había

do, Dr. SHAPIRO (cuya unidad en la Universidad de Boston había recibido más de millón y medio de dólares de Merrell Dow), testificó que un fármaco tomado por una madre después de la formación de las extremidades del feto no puede causar defectos congénitos en las extremidades; pero en el interrogatorio cruzado reconoció que los datos en los que fundaba su opinión agrupaban tanto a mujeres que tomaron Bendectin en el periodo de formación de las extremidades y mujeres que lo tomaron después; y que por tal razón su estudio había sido criticado por artículos posteriores<sup>115</sup>.

El experto de la demandada, Dr. NEWBERNE, vicepresidente de Merrell Dow y responsable de los estudios en animales y la seguridad de los fármacos, testificó que mientras el estudio del Dr. SMITHELLS que concluía que el Bendectin no era teratogénico había sido arbitrado y rechazado por la *British Medical Journal*, el *Lancet* y la *New England Journal of Medicine*, y finalmente aceptado por una revista de mucho menos prestigio, *Teratology*, el autor había estado activamente solicitando fondos a la compañía, escribiendo que esa publicación podía ser de valor para Merrell Dow «[dado que ésta] puede salvar a la compañía de pagar grandes sumas de dinero en el tribunal de California [...]»<sup>116</sup>.

Y el editor de *Teratology*, Dr. R. BRENT<sup>117</sup>, quien había sido contratado durante dieciocho años como perito por Merrell Dow, testificó que su única educación formal en epidemiología había sido un curso en estadísticas, pero que él se consideraba a sí mismo como una autoridad internacional en «*secular trend data*» —un área en la cual, como comentó el juez BERNSTEIN, aparentemente él era el único practicante<sup>118</sup>. Usando su prerrogativa editorial para esquivar el *referato*<sup>119</sup>, había publicado en su propia revista un artículo titulado «Litigation-Produced Pain, Disease, and Suffering: An Experience With Congenital Malformation Lawsuits» («El dolor, la enfermedad y el sufrimiento producidos por los litigios: una experiencia con demandas por malformación congénita»), basado en su revisión de las deposiciones y los expedientes de los juicios, concluyendo que los 17 demandantes de los 17 procesos judiciales mentían<sup>120</sup>; había sometido un borrador de éste artículo, titulado «Bendectin: The Most Comprehensively Studied Human Non-Teratogen, and the Foremost Tortogen-Litigen» («Bendectin: el teratógeno humano más exhaustivamente estudiado y el daño que está en el origen de más procesos judiciales»), a los abogados de Merrell Dow para que lo

---

usado en su artículo sobre cataratas y vómitos; y estaba de acuerdo en que un artículo «menos que bueno» podría pasar una revisión por pares.

<sup>115</sup> *Ibid.*, 215-217.

<sup>116</sup> *Ibid.*, 218-222. El Dr. NEWBERNE también reconoció que después de un primer estudio realizado por el Dr. HENDRICKS encontró un incremento estadístico significativo en defectos en el corazón de monos que habían sido tratados con Bendectin, Merrell Dow financió un segundo estudio que llegó a resultados más favorables para la compañía. La entrada en los registros financieros de Merrell Dow fue: «Estudio de Hendricks con monos-defensa», *ibid.*, 221.

<sup>117</sup> En el caso *Wells v. Ortho Pharm. Corp.*, 615 F. Supp. 262, 281 (N.D. Ga. 1985) (citado en Blum 1996 (*supra* nota 108), 10 (con un voto particular en contra de J. CASTILLE), el testimonio del Dr. BRENT se encontró no-creíble porque «[su] testimonio y forma sugerían un grado de convicción no justificado en sus propias conclusiones en una disciplina en la que [...] las explicaciones son sólo más o menos probables».

<sup>118</sup> Blum 1996, *supra* nota 108, 223-224.

<sup>119</sup> *Vid. ibid.*, 224 (explicando que algunos de los escritos del Dr. BRENT habían sido publicados en *Teratology* solamente «por su prerrogativa editorial personal»).

<sup>120</sup> BRENT, 1997: 5, tabla 1.

editaran, esperando que fuera publicado en el *New England Journal*, en el *Journal of the American Medical Association* o en el *Lancet*<sup>121</sup>.

Finalmente, Merrell Dow se impuso; en el 2000 la Corte Suprema de Pensilvania decidió que aun cuando la prueba pericial presentada por los Blum podría decirse que fue admisible bajo (lo que parece ser) «algo menos exacto que el estándar *Daubert*», es inadmisble bajo el criterio *Frye*<sup>122</sup>. Sin embargo, el magistrado CASTILLE, en un voto particular, recordó a sus colegas que el tribunal de primera instancia, citando a *Frye*, se había dejado influir por el hecho del «consenso científico [sobre el Bendectin], derivado en gran medida de la influencia de la compañía y los intereses judiciales de la contraparte»; y que «gran parte de la “ciencia” en esta área, defendida por Merrell Dow como el punto de vista científico objetivo y aceptado por la generalidad que exige la exclusión de las conclusiones “contrarias” hechas por los “peritos” de los demandantes, era en sí misma un producto de la influencia de Merrell Dow en la resolución del proceso»<sup>123</sup>.

Sin embargo, para los objetivos actuales son las conclusiones del juez BERNSTEIN las más pertinentes. El testimonio en este caso, observa, «demuestra cómo el “consenso científico” puede ser creado a través de la compra de investigación y la manipulación de la literatura “científica”, financiado como parte de la defensa en un proceso y coreografiado por abogados»<sup>124</sup>. Éste «claramente demuestra que no todas las revistas “con *referato*” son creadas iguales», que «no todos los artículos contenidos en revistas “con *referato*” son incluso evaluados...», y que «[a]rtículos son puestos intencionalmente en revistas con *referato* para usarlos ante un tribunal»<sup>125</sup>. Y en el Apéndice B de su opinión, titulada «Science and Justice», el juez BERNSTEIN añade:

El testimonio demostró una literatura médico-científica en revistas con *referato* que es creada y manipulada para su uso ante tribunales. El testimonio demostró que algunos artículos son puestos en revistas «con *referato*», sin la revisión de autoridades independientes y editados por abogados [...] [y] reveló la edición de los hechos, supuestamente resultados de investigación científica, por los mismos abogados del demandado en un proceso<sup>126</sup>.

\* \* \*

El ejemplo es instructivo, refuerza el reconocimiento del magistrado BLACKMUN respecto a que la publicación arbitrada no es una garantía de la «validez científica» sino que, en su lectura más optimista, es un indicador muy falible; y nos recuerda que si los tribunales plantearan las preguntas sugeridas aquí, este factor *Daubert* pudiera, y debería, ser manejado con más cautela y más sutileza de lo que ha sido, en general, hasta ahora. También es muy preocupante; pues sugiere que la literatura científica

<sup>121</sup> *Blum* 1996, *supra* nota 108, 223-228. El artículo en cuestión de hecho había sido publicado, aunque la sentencia del juez BERNSTEIN no nos lo dijese. *Vid.* BRENT, 1995 (la versión publicada del artículo del Dr. BRENT). El artículo dio lugar a una demanda por difamación en contra del Dr. BRENT por el Dr. S. NEWMAN, quien fue uno de los peritos del demandando en *Blum* y a quien el Dr. BRENT citó erróneamente. *Newman v. Brent*, 1998 U.S. Dist. LEXIS 10476 (D.D.C. 1998).

<sup>122</sup> *Blum* 2000, *supra* nota 108, 3. A pesar de la retórica de la Corte en *Daubert*, sin embargo, es cuestionable, por decir lo menos, si el criterio *Daubert* realmente es «menos riguroso» que el criterio *Frye*.

<sup>123</sup> *Ibid.*, 16-20.

<sup>124</sup> *Blum* 1996, *supra* nota 108, 230.

<sup>125</sup> *Ibid.*, 246-247.

<sup>126</sup> *Ibid.*, 248-249.

usada en un proceso judicial, por la que un criterio como «la evaluación por pares y la publicación» entró en el vocabulario jurídico oficial para la admisibilidad, puede haber sido contaminada por los intereses en litigio. Irónicamente, parece que la misma comercialización de la investigación médica que ha contribuido a la corrupción progresiva de la evaluación por pares y la publicación, también puede haber sido en parte responsable de la confianza que el sistema jurídico ha estado teniendo en ese proceso como un factor para determinar la fiabilidad probatoria<sup>127</sup>.

(Traducción de Carmen Vázquez)

## BIBLIOGRAFÍA CITADA

- ALTMAN, L. K., 2006: «For Science's Gatekeepers, a Credibility Gap», en *New York Times*, May 2, F1.
- ARMSTRONG, D., 2006: «Bitter Pill: How the New England Journal of Medicine Missed Warning Signs in Vioxx-Medical Weekly Waited Years to Report Flaws in Article that Praised Pain Drug-Merck Seen as "Punching Bag"», en *Wall Street Journal*, May 15, A1, A10.
- BAILAR, J. C., 1991: «Reliability, Fairness, Objectivity, and Other Inappropriate Goals in Peer Review», en *Behavioral and Brain Sciences*, 14, 137-138.
- BEGLEY, S., 2006: «Science Journals Artfully Try to Boost Their Rankings», en *Wall Street Journal*, June 5, B1, B8.
- BERO, L. *et al.*, 1992: «The Publication of Sponsored Symposia in Medical Journals», en *New England Journal of Medicine*, 327, 1135-1140.
- BOMBADIER, C. *et al.*, 2000: «Comparison of Upper Gastrointestinal Toxicity of Rofecoxib and Naproxen in Patients With Rheumatoid Arthritis», en *New England Journal of Medicine*, 343, núm. 21, 1520-1528.
- BRENT, R. L., 1995: «Bendectin: Review of the Medical Literature of a Comprehensively Studied Human Nonteratogen and the Most Prevalent Tortogen-Litigen», en *Reproductive Toxicology*, 9, núm. 4, 337-349.
- 1997: «Litigation-Produced Pain, Disease, and Suffering: An Experience with Congenital Malformation Lawsuit», en *Teratology*, 16, núm. 1, 1-13.
- BRESALIER, R. S. *et al.*, 2005: «Cardiovascular Events Associated with Rofecoxib in a Colorectal Adenoma Chemoprevention Trial», en *New England Journal of Medicine*, 352, 1092-1102.
- BRYANT, D. *et al.*, 2006: «How Many Patients? How Many Limbs? Analysis of Patients or Limbs in the Orthopedic Literature: A Systematic Review», en *J. Bone & Joint Surgery*, 88, 41.
- BRIDGMAN, P., 1933: «The Struggle for Intellectual Integrity», en P. BRIDGMAN, 1995: *Reflections of a Physicist*, New York: *Philosophical Library*, 2.<sup>a</sup> ed.
- BURNHAM, J. C., 1990: «The Evolution of Editorial Peer Review», en *Journal of the American Medical Association*, 263, 1323-1329.
- CAREY, B., 2006: «Researcher Pulls his Name from Paper on Prayer and Fertility», en *New York Times*, December 4, A15.
- CHAN, A. W. *et al.*, 2004: «Empirical Evidence for Selective Reporting of Outcomes in Randomized Trials», en *Journal of the American Medical Association*, 291, 2457-2465.

<sup>127</sup> Mi agradecimiento a M. MIGOTTI por sus útiles comentarios a dos borradores previos; a L. TILSON y S. SHOTT por su ayuda para encontrar material sobre errores estadísticos y retractaciones en artículos publicados; y a los diversos trabajadores de la biblioteca de derecho de la Universidad de Miami, especialmente a D. HOLLANDER y B. BRANDOM, quienes me ayudaron a encontrar material pertinente.

- CHAN, E. J., 1995: «The “Brave New World” of *Daubert*: True Peer Review, Editorial Peer Review, and Scientific Validity», en *NYU Law Review*, 70, 100-134.
- COUZIN, J., y UNGER, K., 2006: «Cleaning up the Paper Trail», en *Science*, 312, 38-43.
- EDITORIAL, 1983: «Is Science Really a Pack of Lies?», en *Nature*, 303, 361.
- FLAMM, B., 2004: «The Columbia University “Miracle” Study: Flawed and Fraud», en *Skeptical Inquirer*, 28, 25-31.
- FRIEDMAN, P. J., 1990: «Correcting the Literature Following Fraudulent Publication», en *Journal of the American Medical Association*, 263, 1416-1419.
- GARCÍA-BERTHOU, E., y ALCARAZ, C., 2004: «Incongruence between Test Statistics and P Values in Medical Papers», en *BMS Medical Research Methodology*, 4, núm. 13, 1-5.
- GARDNER, M. J., y BOND, J., 1990: «An Exploratory Study of Statistical Assessment of Papers Published in The British Medical Journal», en *Journal of the American Medical Association*, 263, 1355.
- GRAVITZ, L., 2006: «Biology’s Image Problem», en *Rockefeller University Scientist*, 1, núm. 1, 10-12.
- HAACK, S., 2003: *Defending Science-Within Reason: Between Scientism and Cynicism*, Amherst, NY: Prometheus Books.
- 2008: «What’s Wrong with Litigation-Driven Science?», en *Seton Hall Law Review*, 38, núm. 3, 1053-1083. Reimpreso en S. HAACK, 2014: *Evidence Matters: Science, Proof, and Truth in the Law*, New York: Cambridge University Press, 180-207.
- HALLER, Ch., y BENOWITZ, N. L., 2000: «Adverse Cardiovascular and Central Nervous System Events Associated with Dietary Supplements Containing Ephedra Alkaloids», en *New England Journal of Medicine*, 343, 1833-1838.
- HENDERSON, D., 2005: «Journal Says Vioxx Woes Suppressed. Merck Blamed: Correction Sought», en *Boston Globe*, 9 de diciembre.
- HORROBIN, D. F., 1990: «The Philosophical Basis of Peer Review and the Suppression of Innovation», en *Journal of the American Medical Association*, 263, 1438-1441.
- HORTON, R., 2006: «Expression of Concern: Non-Steroidal Anti-Inflammatory Drugs and the Risk of Oral Cancer», en *Lancet*, 367.
- IONNADIS, J. P. A., 2005: «Contradicted and Initially Stronger Effects in Highly Cited Clinical Research», en *Journal of the American Medical Association*, 294, 218-228.
- JUDSON, H. F., 1994: «Structural Transformations of the Sciences and the End of Peer Review», en *Journal of the American Medical Association*, 272, 92-94.
- KRONICK, D. A., 1990: «Peer Review in 18th-Century Scientific Journalism», en *Journal of the American Medical Association*, 263, 1321-1322.
- LOCK, S., 1985: *A Difficult Balance: Editorial Peer Review in Medicine*, London: Nuffield Provincial Hospitals Trust.
- LOCK, S., y SMITH, J., 1990: «What Do Peer Reviewers Do?», en *Journal of the American Medical Association*, 263, 1341-1343.
- MARTINSON, B. C. *et al.*, 2005: «Scientists Behaving Badly», en *Nature*, 435, 737-738.
- MAXWELL, S. R. J., y WEBB, D. J., 2005: «Cox-2 Selective Inhibitors-Important Lessons Learned», en *Lancet*, 365, 449-451.
- MCCUTCHEN, Ch. W., 1991: «Peer Review: Treacherous Servant, Disastrous Master», en *Technology Review*, 94, 28-51.
- NEVILLE, J. A. *et al.*, 2006: «Errors in the Archives of Dermatology and the Journal of the American Academy of Dermatology from January through December 2003», en *Archives Dermatology*, 142, 737-38.

- OSMOND, D. H., 1983: «Malice's Wonderland: Research Funding and Peer Review», en *Journal of Neurobiology*, 14, núm. 2, 95-112.
- PFEIFFER, M. P., y SNODGRASS, G. L., 1990: «The Continued Use of Retracted, Invalid Scientific Literature», en *Journal of the American Medical Association*, 263, 1420-1423.
- POLANYI, M., 1946: *Science, Faith, and Society*, Cambridge: Cambridge University Press.
- POSNER, R. A., 2004: «Against the Law Reviews», en *Legal Affairs*, 57-58.
- RENNIE, D., 1986: «Guarding the Guardians: A Conference on Editorial Peer Review», en *Journal of the American Medical Association*, 256, 2391-2392.
- 2003: «Editorial Peer Review: Its Development and Rationale», en F. GODLEE y T. JEFFERSON (eds.): *Peer Review in Health Sciences*, London: BMJ Publishing Group, 2.<sup>a</sup> ed.
- SMITH, R., 1997: «Peer Review: Reform or Revolution?», en *British Medical Journal*, 315, 759-760.
- SOX, H. C., y RENNIE, D., 2006: «Research Misconduct, Retraction, and Cleansing Medical Literature: Lessons from the Poehlman Case», en *Annals of Internal Medicine*, 144, 609-613.
- SUDBØ, J. et al., 2005: «Non-Steroidal Anti-Inflammatory Drugs and the Risk of Oral Cancer: A Nested Case-Control Study», en *Lancet*, 366, 1359-1366.
- SUN, M., 1989: «Peer Review Comes under Peer Review», en *Science*, 224, 910-912.
- VAIL, A., y GARDENER, E., 2003: «Common Statistical Errors in the Design and Analysis of Subfertility Trials», en *Human Reproduction*, 18, núm. 5, 1000-1004.
- WADE, N., y SANG-HUN, Ch., 2006: «Human Cloning Was All Faked, Koreans Report», en *New York Times*, January 10, A1.
- WELLER, A. C., 1990: «Editorial Peer Review in U.S. Medical Journals», en *Journal of the American Medical Association*, 263, 1344-1345.
- WON TESORIERO, H., 2007: «Vioxx Study Correction May Add Pressure to Merck's Defense», en *Wall Street Journal*, June 27, A2.
- WYSOCKI, B. Jr., 2005: «Scholarly Journals' Premier Status is Diluted by Web», en *Wall Street Journal*, May 23, A1, A8.
- YANKAUER, A., 1990: «Who Are the Peer Reviewers and How Much Do They Review?», en *Journal of the American Medical Association*, 263, 1338-1340.
- ZIMAN, J. M., 1968: *Public Knowledge*, Cambridge: Cambridge University Press.
- 1969: «Information, Communication, Knowledge», en *Nature*, 318-324.
- ZUCKERMAN, H., y MERTON, R. K., 1971: «Patterns of Evaluation in Science: Institutionalism, Structure, and Functions of the Referee System», en *Minerva*, 9, 66-100.